

JP10222513A: MULTILINGUAL CORRESPONDENCE COMMUNICATION SYSTEM

[View Images \(1 pages\)](#)

Inventor(s): **KURACHI ATSUSHIKO
FUJII NORIHISA
KAWAI ATSUSHI**

Applicant(s): **BROTHER IND LTD**

Issued/Filed Dates: **Aug. 21, 1998 / Jan. 31, 1997**

Application Number: **JP1997000018689**

IPC Class: **G06F 017/28; G06F 013/00;**

Abstract: **Problem to be solved:** To provide a multilingual correspondence communication system which does not need a special language translation program for a client device and a server device and also more easily offers document data that is described in many languages.

*uses an intermediate
language for translation
to any language*

Solution: A server device 102 stores intermediate format document data which is acquired as a result of semantically and grammatically analyzing document data that is described in a natural language. When a client device 103 makes a request for document data, intermediate format document data that corresponds to the document data is sent from the device 102 to a converter 104. The converter 104 generates document data that is described in a natural language which is requested by the device 103 from the received intermediate format document data and sends the generated document data to the device 103.

COPYRIGHT: (C)1998,JPO

Foreign References: **none**

(No patents reference this one)



**Nominate this
invention
for the Gallery...**

**Alternate
Searches**



[Patent Number](#)



[Boolean Text](#)



[Advanced Text](#)



**SEARCH PATENT FULL TEXT
WITH NATURAL LANGUAGE**

-6- (WPAT)

AN - 98-511448/44

XRPX- N98-399161

TI - Multilanguage correspondence communication system for internet - converts intermediate format document data generated from natural language into predetermined language and transmits it to client

DC - T01

PA - (BRER) BROTHER KOGYO KK

PR - 97.01.31 97JP-018689

NUM - 1 patent(s) 1 country(s)

PN -- JP10222513 A 98.08.21 * (9844) 19p G06F-017/28

AP -- 97JP-018689 97.01.31

IC1 - G06F-017/28

IC2 - G06F-013/00

AB - JP10222513 A

The system makes use of a client (103) which transmits the document identification information that specifies a document data to a converter (104) and a server (102) using two demand information transmission units. An intermediate format document data corresponding to the document identification information is generated by the server by gramatic and semantic analysis of a natural language and is transmitted to the convertor.

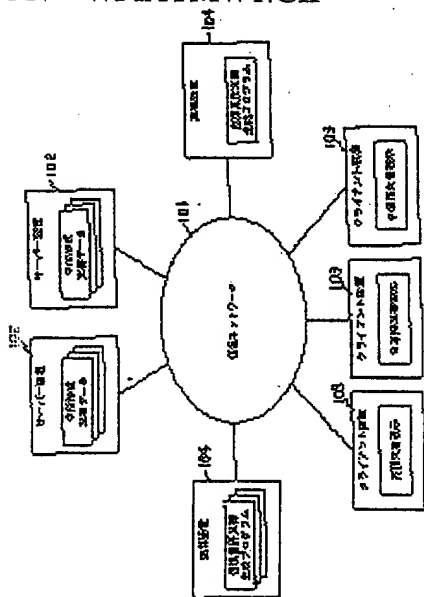
The convertor then generates the document data indicating the desired language of the client from the intermediate format document data. Then, the document data generated by the convertor is transmitted to the client.

ADVANTAGE - Avoids need for special language translation program. Offers document data indicating desired language. Increases data processing speed.

(Dwg.1/14)

MC - EPI: T01-H07C3C T01-J14

FN - WPIAYMW1.GIF



(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平10-222513

(43) 公開日 平成10年(1998) 8月21日

(51) Int.Cl.⁵

識別記号

F I

G 0 6 F 17/28

G 0 6 F 15/38

Z

13/00

3 5 7

13/00

3 5 7 Z

15/38

C

審査請求 未請求 請求項の数 5 O L (全 19 頁)

(21) 出願番号

特願平9-18689

(22) 出願日

平成9年(1997) 1月31日

(71) 出願人 000005267

ブラザー工業株式会社

愛知県名古屋市瑞穂区苗代町15番1号

(72) 発明者

倉知 敦彦

名古屋市瑞穂区苗代町15番1号ブラザー工業株式会社内

(72) 発明者

藤井 則久

名古屋市瑞穂区苗代町15番1号ブラザー工業株式会社内

(72) 発明者

河合 淳

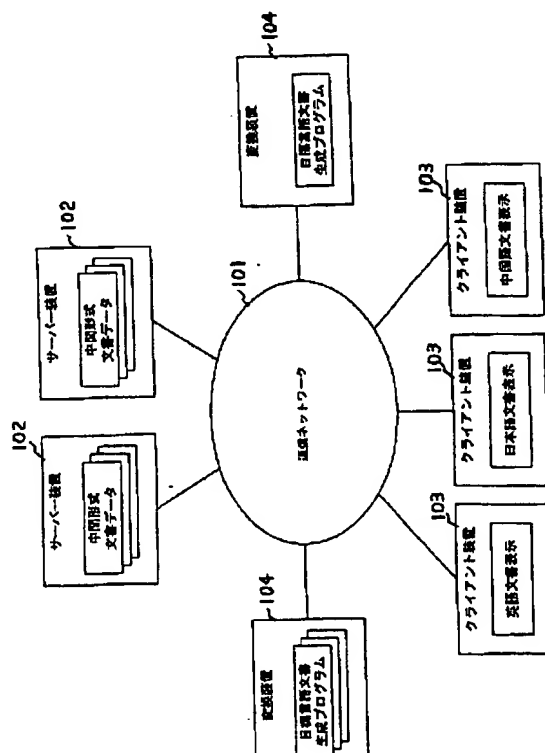
名古屋市瑞穂区苗代町15番1号ブラザー工業株式会社内

(54) 【発明の名称】 多言語対応通信システム

(57) 【要約】

【課題】 クライアント装置やサーバー装置に特別な言語翻訳プログラムを必要とせず、かつ多数の言語で記述される文書データをより簡易に提供することができる多言語対応通信システムを提供することである。

【解決手段】 サーバー装置102に自然言語で記述された文書データを文法的、意味的に解析した結果として得られる中間形式文書データを記憶する。クライアント装置103から文書データの要求があると、その文書データに対応した中間形式文書データをサーバー装置102から変換装置104に送信する。変換装置104は受信した中間形式文書データからクライアント装置103が要求する自然言語で記述された文書データを生成し、この生成された文書データをクライアント装置103に送信する。



【特許請求の範囲】

【請求項1】 クライアント装置と、前記クライアント装置からの要求に応じた文書データを前記クライアント装置に送信するサーバー装置と、前記クライアント装置と前記サーバー装置との通信を変換し、必要に応じて文書データを変換する変換装置とから構成される多言語対応通信システムであって、前記クライアント装置は、前記文書データを特定可能な文書識別情報を前記変換装置に送信する第1の要求情報送信手段と、前記変換装置より前記文書データを受信し、出力する受信出力手段とを備え、前記変換装置は、前記第1の要求情報送信手段により送信された情報を受信する第1の要求情報受信手段と、前記文書識別情報を前記サーバー装置に送信する第2の要求情報送信手段と、自然言語で記述された文書を文法的、意味的に解析した結果として得られる中間形式表現で記述した中間形式文書データを、前記サーバー装置から受信する第1の文書受信手段と、前記第1の文書受信手段により受信した中間形式文書データから、所定の言語で記述される文書データを生成する目標言語文書生成手段と、前記目標言語文書生成手段で生成された文書データを、前記クライアント装置に送信する第1の文書送信手段とを備え、前記サーバー装置は、前記第2の要求情報送信手段により送信された情報を受信する第2の要求情報受信手段と、少なくとも一つの文書データを、前記中間形式文書データとして記憶する中間形式文書記憶手段と、前記文書識別情報に対応した中間形式文書データを、前記変換装置に送信する第2の文書送信手段とを備えたことを特徴とする多言語対応通信システム。

【請求項2】 前記変換装置は、

所定の言語で記述された文書データである原文書データを前記サーバー装置から受信する第2の文書受信手段と、前記第2の文書受信手段で受信した原文書データを解析し、その原文書データの解析結果から中間形式文書データを生成する原文書解析手段と、前記原文書解析手段により生成された中間形式文書データを前記サーバー装置に送信する第3の文書送信手段とをさらに備え、前記サーバー装置は、原文書データを前記変換装置に送信する第4の文書送信手段と、前記第3の文書送信手段により送信された中間形式文書データを受信する第3の文書受信手段とをさらに備え、前記中間形式文書記憶手段は前記第3の文書受信手段で

受信された中間形式文書データを記憶することを特徴とする請求項1に記載の多言語対応通信システム。

【請求項3】 前記変換装置は、前記第1の文書受信手段で受信した文書が中間形式表現で記述した中間形式文書データであるか、あるいは自然言語で記述した文書データであるかの判定を行う文書表現判定手段をさらに備え、

前記第1の文書受信手段は、中間形式文書データに加えて、自然言語で記述された文書データも受信し、

10 前記第1の文書送信手段は、前記文書表現判定手段により、自然言語で記述された文書データであると判定された場合は、前記第1の文書受信手段で受信した文書データを前記クライアント装置に送信し、中間形式文書データであると判定された場合は、前記目標言語文書生成手段で生成された文書データを前記クライアント装置に送信することを特徴とする請求項1または2に記載の多言語対応通信システム。

【請求項4】 前記第1の要求情報送信手段は、前記文書識別情報に加えて、前記文書データの記述言語を指定する要求言語情報も送信し、

20 前記目標言語文書生成手段は、前記要求言語情報に対応した言語で記述した文書データを生成することを特徴とする請求項1乃至3のいずれかに記載の多言語対応通信システム。

【請求項5】 前記サーバー装置は、中間形式文書データの元となった自然言語で記述された文書データである原文書データを、中間形式文書データと対応づけて記憶する原文書記憶手段と、

30 前記原文書記憶手段で記憶され、受信した前記文書識別情報に対応した原文書データの記述言語が、受信した要求言語情報に対応した言語であるか否かを判定する原文書判定手段とをさらに備え、

前記第2の要求情報送信手段は、前記文書識別情報に加えて、前記要求言語情報も送信し、

40 前記第2の文書送信手段は、前記原文書判定手段により、前記要求言語情報に対応した言語であると判定された場合は、前記原文書記憶手段で記憶され、受信した前記文書識別情報に対応した原文書データを前記変換装置に送信し、前記要求言語情報に対応した言語ではないと判定された場合は、前記文書識別情報に対応した中間形式文書データを、前記変換装置に送信することを特徴とする請求項2乃至4のいずれかに記載の多言語対応通信システム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、翻訳機能を持つ変換装置と、文書データを提供することが可能な文書データのサーバー装置と、文書データの表示が可能なクライアント装置とを通信ネットワークを介して接続した多言語対応通信システムに関するものである。

【0002】

【従来の技術】従来、インターネットに代表されるように、通信ネットワークを介して、文書データを検索、閲覧するシステムが普及している。この様なシステムは、各種の文書データを記憶する少なくとも1つのサーバー装置と、文書データの表示が可能な少なくとも1つのクライアント装置と、それらを接続する通信ネットワークとで構成される。通信ネットワークは、公衆回線網や、専用回線網あるいはその組み合わせから構成される。サーバー装置は、各種の文書データが記憶できる記憶装置と、通信ネットワークへの接続装置とを備え、クライアント装置は文書データを表示可能な表示装置と、閲覧したい文書の指示情報を入力できる入力装置と、通信ネットワークへの接続装置とを備える。

【0003】例えば、インターネットでは、サーバー装置、クライアント装置共に非常に多くの装置が接続されており、その所在地も全世界に広がっている。そして、それらを接続する通信ネットワークを介して世界各地の様々な情報を、検索、閲覧することが可能なシステムとなっている。

【0004】このようなインターネットでは図14に示すように、各サーバー装置1401には、英語とか、日本語、中国語等の様々な言語で記述された文書データが記憶されている。通信ネットワーク101を介してサーバー装置に接続されているクライアント装置1402では、これらの様々な言語で記述された文書データをそれぞれの言語で表示することが可能である。また、言語翻訳プログラムを内蔵するクライアント装置もあり、必要に応じて所望の言語に翻訳して文書データを表示することも可能である。

【0005】

【発明が解決しようとする課題】上述したシステム、特にインターネットのように全世界に広がったネットワークを介して様々な文書データを閲覧する場合、文書データの提供者と、その文書データの閲覧者とは異なる言語を母国語としていることが多い。しかし母国語以外の言語での記述や、その理解には困難を伴うことが多い。これを解決するため、次の2つの方法が知られている。1つは、同一内容の文書を複数の言語で記述された複数の文書データとしてサーバー装置に記憶させ、所望の言語の文書データをクライアント装置から選択することで、所望の言語で記述された文書データを閲覧可能とする方法である。他の方法は、言語翻訳プログラムをクライアント装置に搭載し、その言語翻訳プログラムを用いてサーバーから送信されてきた文書データを所望する言語の文書データに翻訳する方法である。これらの方法により、母国語以外の言語で記述された文書データも母国語で閲覧可能となる。しかし、その場合でも以下のような問題が残されている。

【0006】先ず、同一内容の文書を複数の言語で記述

された複数の文書データとしてサーバー装置に記憶させる方法では、1つの内容の文書でも、その内容を様々な言語で記述された複数の文書データとして作成する必要があり、多くの費用と労力を必要とする問題がある。

【0007】また、言語翻訳プログラムをクライアント装置に搭載する方法では、各サーバーにより提供される様々な言語の文章を全て翻訳して表示するためには、クライアント装置にそれぞれの言語から所望の言語へ翻訳可能な多数の言語翻訳プログラムを搭載する必要がある。これはクライアント装置のコストを上昇させる。また、現在実用化されている言語翻訳プログラムでは、翻訳する原文書に元来含まれている曖昧さにより、望ましい翻訳結果が得られないと言う問題もある。

【0008】これを解決するために、本出願人が既に出願した特願平8-322842号には、サーバー装置に記憶される文書データを、自然言語で記述された文書データではなく、自然言語で記述された文書データを文法的、意味的に解析した結果として得られる中間形式文書データとして記憶する。そして、クライアント装置から文書データの要求があると、その文書データに対応した中間形式文書データを記憶装置から読み出し、読み出した中間形式文書データからクライアント装置が要求する自然言語で記述された文書データを生成し、生成された文書データをクライアント装置に送信する。このように、クライアント装置に特別な言語翻訳プログラムを必要とせず、かつ多数の言語で記述される文書データをより簡易に提供可能な多言語対応通信システムが記述されている。

【0009】しかし、その既提案の多言語対応通信システムにおいても、次のような問題点が残されている。

【0010】サーバー装置には、中間形式文書データから自然言語で記述された文書データを生成するプログラムを、多数の言語に対応できるように持たなければならないが、これはサーバー装置のコストを上昇させる。また、中間形式文書データから自然言語で記述された文書データを生成するためには、生成する言語に対応した各種の辞書データが必要とされる。このような辞書データは、必要に応じ追加、修正等の保守が適宜行われないと、望ましい翻訳結果が得られない場合があるため、各サーバー装置毎に辞書データの管理、保守を行う必要があるが、このための費用や労力が必要となる。

【0011】本発明は、上述した問題点を解決するためになされたものであり、クライアント装置に特別な言語翻訳プログラムを必要とせず、かつ多数の言語で記述される文書データをより簡易に提供可能であって、さらに、サーバー装置を簡易に構成でき、辞書の管理も簡易に行うことが可能な多言語対応通信システムを提供することを目的とする。

【0012】

【課題を解決するための手段】この目的を達成するため

に、本発明の請求項1に記載の多言語対応通信システムは、クライアント装置と、前記クライアント装置からの要求に応じた文書データを前記クライアント装置に送信するサーバー装置と、前記クライアント装置と前記サーバー装置との通信を変換し、必要に応じて文書データを変換する変換装置とから構成される。前記クライアント装置では、第1の要求情報送信手段が前記文書データを特定可能な文書識別情報を前記変換装置に送信し、受信出力手段が前記変換装置より前記文書データを受信し、表示出力、印字出力、音声出力等の出力を行う。また、前記変換装置では、第1の要求情報受信手段が前記第1の要求情報送信手段により前記クライアント装置から送信された情報を受信し、第2の要求情報送信手段が前記文書識別情報を前記サーバー装置に送信し、第1の文書受信手段が自然言語で記述された文書を文法的、意味的に解析した結果として得られる中間形式表現で記述した中間形式文書データを前記サーバー装置から受信する。

【0013】そして、目標言語文書生成手段が前記第1文書受信手段により受信した中間形式文書データから、所定の言語で記述される文書データを生成し、第1の文書送信手段が前記目標言語文書生成手段で生成された文書データを、前記クライアント装置に送信する。前記サーバー装置では、第2の要求情報受信手段が前記第2の要求情報送信手段により前記変換装置から送信された情報を受信し、中間形式文書記憶手段が、少なくとも一つの文書データを前記中間形式文書データとして記憶し、第2の文書送信手段が前記文書識別情報に対応した中間形式文書データを前記変換装置に送信する。

【0014】このように、クライアントから所望の記述言語を指定でき、それに従って中間形式文書データの形式で記憶された文書データから、変換装置が自然言語で記述された文書データを生成し、クライアント装置で出力できるので、サーバー装置に多数の言語で記述した文書データを保存したり、クライアント装置やサーバー装置に言語翻訳プログラムを搭載することなく簡易に構成でき、所望の言語で表記された文書データの閲覧ができる。

【0015】また、請求項2に記載の多言語対応通信システムでは、前記変換装置において、第2の文書受信手段が所定の言語で記述された文書データである原文書データを前記サーバー装置から受信し、原文書解析手段が前記第2の文書受信手段で受信した原文書データを解析し、そしてその原文書データの解析結果から中間形式文書データを生成し、第3の文書送信手段が前記原文書解析手段により生成された中間形式文書データを前記サーバー装置に送信する。前記サーバー装置においては、第4の文書送信手段が原文書データを前記変換装置に送信し、第3の文書受信手段が前記第3の文書送信手段により前記変換装置から送信された中間形式文書データを受信し、前記中間形式記憶手段が前記第3の文書受信手段

で受信された中間形式文書データを記憶する。

【0016】このように、自然言語で記述された原文書データを、サーバー装置から変換装置に送信し、変換装置で原文書データから中間形式文書データを生成し、生成された中間形式文書データをサーバー装置に送り返し、サーバー装置でその中間形式文書データを記憶することで、中間言語の作成に要するプログラムやデータをサーバー装置毎に用意する必要がなくなる。

【0017】また、請求項3に記載の多言語対応通信システムでは、前記変換装置において、第1の文書受信手段は、中間形式文書データに加えて自然言語で記述された文書データもサーバー装置から受信し、文書表現判定手段が、その受信した文書が中間形式表現で記述した中間形式文書データであるか、あるいは自然言語で記述した文書データであるかの判定を行う。そして、第1の文書送信手段は、前記文書表現判定手段により前記第1の文書受信手段で受信した文書データが自然言語で記述された文書データであると判定された場合には、前記第1の文書受信手段で受信した文書データを前記クライアント装置に送信し、一方、前記第1の文書受信手段で受信した文書データが中間形式文書データであると判定された場合は、前記目標言語文書生成手段で生成された文書データを前記クライアント装置に送信する。

【0018】これにより、例えば、従来用いられているサーバー装置のような、中間形式文書データが保存されていないサーバー装置や、自然言語で記述された文書データを出力するサーバー装置等に対応することができ、文書データが得られないという可能性を減少させることが出来る。

【0019】また、請求項4に記載の多言語対応通信システムでは、前記第1の要求情報送信手段は、前記文書識別情報に加えて、前記文書データの記述言語を指定する要求言語情報も送信し、前記目標言語文書生成手段は、前記要求言語情報に対応した言語で記述した文書データを生成する。

【0020】これにより、変換装置において複数の言語に対応する文書データを生成することができ、各言語毎に異なった変換装置を用意する必要がなくなる。

【0021】また、請求項5に記載の多言語対応通信システムでは、変換装置において、前記第2の要求情報送信手段は、前記文書識別情報に加えて、前記要求言語情報も送信する。前記サーバー装置において、原文書記憶手段が、中間形式文書データの元となった自然言語で記述された文書データである原文書データを、中間形式文書データと対応づけて記憶し、原文書判定手段が、前記原文書記憶手段で記憶され前記文書識別情報に対応した原文書データの記述言語が、前記要求言語情報に対応した言語であるか否かを判定する。そして、前記第2の文書送信手段は、前記原文書判定手段により、前記原文書記憶手段で記憶され前記文書識別情報に対応した原文書

データの記述言語が、前記要求言語情報に対応した言語であると判定された場合は、前記原文書記憶手段で記憶され前記文書識別情報に対応した原文書データを前記変換装置に送信し、前記原文書記憶手段で記憶され前記文書識別情報に対応した文書データの記述言語が、前記要求言語情報に対応した言語ではないと判定された場合は、前記文書識別情報に対応した中間形式文書データを、前記変換装置に送信する。

【0022】これにより、所望の言語による文書データがサーバー装置に保存されている場合には、その保存されている文書データを、変換装置を経由してクライアント装置に送信することができる。よって、必要な場合のみ中間形式文書データから所望の言語で記述された文書データを行えばよく、処理の高速化を行うことが出来る。

【0023】

【発明の実施の形態】以下に、本発明の多言語対応通信システムを具体化した実施の形態について図面を参照して説明する。

【0024】図1に本実施の形態の多言語対応通信システムの概念図を示す。サーバー装置102は、文書データを、特定の自然言語によらない中間形式文書データ(詳細は後述)として記憶している。変換装置104は、中間形式文書データから目標となる自然言語(以下、目標言語と呼ぶ)で記述された文書データを生成するための目標言語文書生成プログラムを記憶している。クライアント装置103は通信ネットワーク101を介して、文書データの指示情報と、必要によりその文書データの記述言語の指示情報とからなる要求情報を所望の変換装置104に送信する。変換装置104は、クライアント装置103から送信された要求情報に基づいて、指定された中間形式文書データから指定された自然言語による文書データを生成し、生成した文書データを通信ネットワーク101を介してクライアント装置103に送信する。そして、クライアント装置はその文書データを受信し、表示出力、印字出力、音声出力等の出力を行う。

【0025】図2に本発明を好適に適用した実施の形態のサーバー装置のブロック図を示している。

【0026】サーバー装置102は、ROM202と、RAM203、CPU204、表示制御部205、入力部206、通信制御部207及びそれらを接続するバス209により構成される。CPU204は記憶されている各種のプログラムに基づいて様々な制御を実行し、本サーバー装置102の動作の主体をなしている。表示制御部205は表示データの蓄積とディスプレイモニターの制御を行い、モニター208に表示出力を行う。モニター208はCRTディスプレイや、LCDディスプレイ等様々な表示装置を利用することができる。入力部206は、記憶される文書データの入力や、サーバー装置102への指示を入力するためのものであり、キーボード装置である。通信制御部207は、通信網210と接続される。通

信網210は、公衆回線や、専用回線の回線網であり、通信ネットワーク101の構成要素である。通信制御部207は、通信網210とサーバー装置102とのデータ通信を制御し、データの送受信を行う。通信制御部207は、公衆回線や専用回線と接続するためのモデム装置である。

【0027】なお、入力部206は、前記の構成以外に、マウス装置等の指示装置や、音声入力装置、タッチパネル装置等の様々な入力装置や、その組み合わせで構成してもよい。また、通信制御部207は、無線の公衆回線や専用回線と接続するための送受信装置とか、LAN(Local Area Network)に接続するためのイーサネット接続装置、トークンリング接続装置とか、ATM(Asynchronous Transfer Mode)回線網への接続装置等、様々な接続装置でもよい。

【0028】ROM202は各種プログラム等を記憶している。ROM202内にはさらに各種の記憶部が形成されている。通信制御プログラム記憶部202aは、通信制御部207を介した各種通信の制御や、変換装置からの文書データ要求の受信や、生成された文書データの変換装置への送信処理等を行う通信制御プログラムを記憶している。通信制御プログラムと、通信制御部207とが請求項記載の第2の要求情報受信手段と、第2の文書送信手段との主体をなす。文書管理プログラム記憶部202bは、中間形式文書データ記憶部203dに対して中間形式文書データを保存、読み出し、あるいは、検索を行うときの主体となる文書管理プログラムを記憶する。

【0029】なお、ROM202は、ROMに限らず、磁気ディスク装置や、光磁気ディスク装置、あるいはCD-ROM記憶装置等の記憶装置であってもよい。

【0030】RAM203は、中間形式文書データ等のデータの保存や、様々な一時データを記憶するために使用され、次のような各記憶部が内部に形成されている。

【0031】ワークメモリ203aは各プログラムが動作時に必要とする一時的な記憶領域を提供する。送信文書メモリ203bは、中間形式文書データ記憶部203dに記憶された中間形式文書データかあるいは原文書メモリ203cに記憶された原文書データの何れかを変換装置104に送信するために、文書データを一時的に記憶する。原文書メモリ203cは、中間形式文書データの元となる、所定の自然言語で記述された文書データである原文書データを記憶する。原文書データは入力部206や、通信制御部207を介して入力される。原文書メモリ203cが本発明の原文書記憶手段の主体をなす。

【0032】中間形式文書データ記憶部203dは、中間形式文書データを少なくとも1つ記憶している。中間形式文書データ記憶部203dが、本発明の中間形式文書記憶手段の主体をなす。

【0033】なお、RAM203は、RAMに限らず、磁気ディスク装置や、光磁気ディスク装置等の記憶装置であってもよい。また、中間形式文書データ記憶部203dは、ROM202

上に形成してもよい。また、原文書を保存する必要がない場合においては、原文書メモリ203cはなくてもよい。

【0034】図3に本発明を好適に適用した変換装置104のブロック図を示す。

【0035】変換装置104は、ROM302と、RAM303、CPU304、表示制御部305、入力部306、通信制御部307及びそれらを接続するバス309より構成される。CPU304は記憶されている各種のプログラムに基づいて様々な制御を実行し、本変換装置104の動作の主体をなしている。表示制御部305は表示データの蓄積とディスプレイモニターの制御を行い、モニター308に表示出力を行う。モニター308はCRTディスプレイや、LCDディスプレイ等様々な表示装置を利用できる。入力部306は、変換装置104への各種指示を入力するためのものであり、キーボード装置である。通信制御部307は、通信網210と接続される。通信制御部307は、通信網210と変換装置104とのデータ通信を制御し、データの送受信を行う。通信制御部307は、公衆回線や専用回線と接続するためのモデム装置である。

【0036】なお、入力部306は、上記の構成以外に、マウス装置等の指示装置や、音声入力装置、タッチパネル装置等の様々な入力装置や、その組み合わせで構成してもよい。また、通信制御部307は、無線の公衆回線や専用回線と接続するための送受信装置とか、LANに接続するためのイーサネット接続装置、トークンリング接続装置とか、ATM回線網への接続装置等、様々な接続装置でもよい。

【0037】ROM302は各種プログラムや、翻訳辞書のデータ等を記憶している。

【0038】ROM302内にはさらに各種の記憶部が形成されている。通信制御プログラム記憶部302aは、通信制御部307を介した各種通信の制御や、クライアント装置103からの文書データ要求の受信、生成された文書データのクライアント装置への送信、サーバー装置からの文書データの受信、サーバー装置への文書データ要求の送信等の送受信処理を行う通信制御プログラムを記憶している。通信制御プログラムと、通信制御部307とが、本発明の第1の要求情報受信手段と、第2の要求情報送信手段と、第1の文書受信手段と、第1の文書送信手段との主体をなす。

【0039】また、目標言語文書生成プログラム記憶部302bは、中間形式文書データから、各種の目標言語文書データを生成するための、目標言語文書生成プログラムを記憶する。なお、目標言語文書生成プログラムが請求項記載の目標言語文書生成手段の主体をなす。本実施の形態では、生成する文書の目標言語の種類に応じて目標言語文書生成プログラムが別になっているものとし、目標言語文書生成プログラム記憶部302bは、生成する文書の目標言語に応じた少なくとも1つ以上の目標言語文書生成プログラムを記憶する。なお、前記以外に、各目標言語に対応した複数の目標言語文書生成プログラムを田

いるのではなく、全ての目標言語に対応できる単一の目標言語文書生成プログラムを用い、目標言語文書生成プログラム記憶部302bはその単一の目標言語文書生成プログラムを記憶してもよい。あるいは、言語構造として類似した複数の目標言語に対応可能な目標言語文書生成プログラムを用い、それを記憶してもよい。

【0040】中間形式文書データ作成プログラム記憶部302cは、サーバー装置から受信した所定の自然言語で記述された原文書データに基づいて、中間形式文書データを作成する中間形式文書データ作成プログラムを記憶する。なお、中間形式文書データ作成プログラムが本発明の原文書解析手段の主体をなす。

【0041】翻訳辞書記憶部302dは、目標言語文書生成プログラムが処理を行う場合に必要とされる、目標言語に対応した生成規則データ、形態素生成辞書等の各種データを記憶している。さらに、翻訳辞書記憶部302dは、中間形式文書データ作成プログラムが処理を行う場合に必要とされる形態素解析辞書、構文規則データ、構文解析辞書、依存規則データ等の各種データも記憶している。

【0042】なお、ROM302は、ROMに限らず、磁気ディスク装置や、光磁気ディスク装置、あるいはCD-ROM記憶装置等の記憶装置であってもよい。

【0043】RAM303は、様々なデータを記憶するために使用され、次のような各記憶部が内部に形成されている。

【0044】ワークメモリ303aは各プログラムが動作時に必要とする一時的な記憶領域を提供する。受信文書メモリ303bは、サーバー装置から受信した文書データを一時的に記憶する。送信文書メモリ303cは、サーバー装置から受信した中間形式文書データに基づいて、目標言語文書生成プログラム記憶部302bに記憶された目標言語文書生成プログラムが生成した目標言語文書データを、クライアント装置への送信のために一時的に記憶する。

【0045】なお、RAM303は、RAMに限らず、磁気ディスク装置や、光磁気ディスク装置等の記憶装置であってもよい。また、データを表示する必要がなければ、表示制御部305や、モニター308を持たない構成としてもよい。さらに、指示データ等は通信制御部307を介して入力する構成にすれば、入力部306を持たない構成とすることもできる。そして、また、請求項1に記載の多言語対応通信システムに対応した変換装置であれば、中間形式文書データ作成プログラム記憶部を持たない構成としてもよい。

【0046】次に、図4に本発明を好適に適用した実施の形態のクライアント装置のブロック図を示す。

【0047】クライアント装置103は、ROM402と、RAM403、CPU404、表示制御部405、入力部406、通信制御部407及びそれらを接続するバス409より構成される。CPU404は記憶されている各種のプログラムに基づいて様々な制

御を実行し、本クライアント装置103の動作の主体をなしている。表示制御部405は表示データの蓄積とディスプレイモニターの制御を行い、モニター408に表示出力を行う。モニター408はCRTディスプレイや、LCDディスプレイ等様々な表示装置を利用できる。入力部406は、文書データの文書識別情報や、要求言語情報等を入力したり、クライアント装置103への各種指示を入力するためのものであり、キーボード装置である。通信制御部407は、通信網210と接続され、通信網210とクライアント装置103とのデータ通信を制御し、データの送受信を行う。通信制御部407は公衆回線や、専用回線と接続するためのモデム装置である。

【0048】なお、入力部406は、上記の構成以外に、マウス装置等の指示装置や、音声入力装置、タッチパネル装置等の様々な入力装置や、その組み合わせで構成してもよい。また、通信制御部407は、無線の公衆回線や専用回線と接続するための送受信装置とか、LANに接続するためのイーサネット接続装置、トークンリング接続装置とか、ATM回線網への接続装置等、様々な接続装置でもよい。また、表示制御部405とモニター408とを用いる以外に、印字出力装置や、音声出力装置を用いる構成とし、文書データを印字出力や音声出力する構成としてもよい。

【0049】ROM402は各種プログラムを記憶している。

【0050】ROM402内にはさらに各種の記憶部が形成されている。通信制御プログラム記憶部402aは、通信制御部407を介した各種通信の制御や、変換装置104へのデータ送受信処理等を行う通信制御プログラムを記憶している。この通信制御プログラムと、通信制御部407とが本発明の第1の要求情報送信手段の主体をなす。表示プログラム記憶部402bは、変換装置104から受信した文書データを表示制御部405を用いて表示するための表示プログラムを記憶している。表示プログラムと、表示制御部405と、通信制御プログラムと、通信制御部407とが本発明の受信表示手段の主体をなす。要求情報入力プログラム記憶部402cは、ユーザーが入力部406を用いて、所望する文書データの文書識別情報と、中継装置を特定するためのアドレス情報と、必要により要求言語情報とを入力する処理を制御するための、要求情報入力プログラムを記憶する。

【0051】なお、ROM402は、ROMに限らず、磁気ディスク装置や、光磁気ディスク装置、あるいはCD-ROM記憶装置等の記憶装置であってもよい。

【0052】RAM403は、様々な一時データを保存するために使用され、次のような各記憶部が内部に形成されている。ワークメモリ403aは各プログラムが動作時に必要とする一時的な記憶領域を提供する。受信文書メモリ403bは、変換装置104から受信した文書データを一時保存する。なお、RAM403は、RAMに限らず、磁気ディスク装置や、光磁気ディスク装置等の記憶装置であってもよい。

い。

【0053】次に、クライアント装置103の基本的な処理の流れを図5を用いて説明する。まず、クライアント装置103では、所望する文書データの指示情報である要求情報の入力処理を行う(S501)。次に、入力した指示情報からその文書データが所在するサーバー装置を通信ネットワーク上で一意に決定する情報である、サーバー装置のアドレスを取得する(S502)。本実施の形態では、サーバー装置のアドレスとしてTCP/IP(Transmission Control Protocol/Internet Protocol)で規定されるホスト名あるいはIPアドレスを用いるものとするが、他に通信ネットワーク上のアドレスを決定できる情報であれば何でもよい。次に、所望する文書データ指示情報から、その文書データのサーバー装置内での指示情報である文書識別情報を決定する(S503)。次に、所望する文書データがどの自然言語によって記述されるかを指定する要求言語の情報を決定する(S504)。

【0054】次に、変換装置を通信ネットワーク上で一意に決定する情報である、変換装置のアドレスを取得する(S505)。本実施の形態では、変換装置のアドレスもサーバー装置のアドレスと同様に、ホスト名あるいはIPアドレスを用いるものとする。また本実施の形態では、変換装置のアドレスは要求言語と対応させて予め記憶されているものとし、S505ではS504にて決定された要求言語に基づいて変換装置のアドレスを取得するものとする。なお、変換装置のアドレスは、上記の方法以外に、予めクライアント装置に変換装置のアドレスとして1つだけ設定されているアドレスを用いてもよい。あるいは、変換装置のアドレスはサーバー装置のアドレスの一部と対応づけて記憶されているものとし、S502にて決定されたサーバー装置のアドレスに基づいて変換装置のアドレスを取得してもよい。

【0055】次に、決定されたサーバー装置のアドレスと、要求文書の識別情報と、要求言語情報とを、変換装置に送信する(S506)。なお、要求言語情報は変換装置に送信しない構成としてもよい。本実施の形態では、変換装置への送信方法として、TCP/IP上のHTTP(Hyper Text Transfer Protocol)を用いることとする。具体的にはクライアント装置は、図6(a)に示す形式のデータを変換装置に送信する。"GET"はデータ要求を意味し、<ローケーション情報>にはS502で決定されたサーバー装置のアドレスと、S503で決定された文書識別情報とを所定の形式で組み合わせた情報を指定する。<バージョン情報>はHTTPのバージョン情報を指定し、<CRLF>は復帰改行文字を示す。<各種ヘッダー情報>にはHTTPで規定されたヘッダー情報を必要に応じて指定する。"Accept-Language:"は次に要求言語情報が続くことを示す文字列であり、<要求言語>には要求言語を識別する予め定められた文字列を指定する。

【0056】図6(b)はサーバー装置のアドレスである

ホスト名が“server1”、文書識別情報として“/pub/sample.html”、HTTPバージョンが1.0、日本語を要求言語とした場合の例である。なお、上述したTCP/IPおよびHTTPは、すでに広く実用化された公知の技術であるため、説明を省略する。

【0057】S506の後、変換装置から送信される要求文書のデータやエラー情報等の情報を受信し(S507)、受信したデータを表示制御部405を用いてモニター408等に表示出力し(S508)、クライアント装置の処理を終了する。

【0058】次に、変換装置104の基本的な処理の流れを図7を用いて説明する。先ず、クライアント装置103がS506にて送信した、サーバー装置のアドレス、要求文書の識別情報、要求言語情報等をクライアント装置から受信する(S701)。次に、S701にて受信したサーバー装置に要求文書の識別情報と、要求言語情報とを送信する(S702)。本実施の形態では、S702にて送信するデータはS506にて説明したHTTPを用いるとする。例示するならば、クライアント装置から図6(b)を受信したとすると、そこからサーバー装置のアドレス情報を削除した図6(c)に表されるような情報となる。次に、S702で送信した情報に対応した文書データを、サーバー装置から受信し、受信文書メモリ303bに記憶する(S703)。

【0059】そして、サーバー装置から受信した文書データが中間形式文書データか否かを判定し(S704)、中間形式文書データである場合はS705に、自然言語で記述された文書データである場合はS707に処理を移す。S705では、受信した要求言語情報に対応した目標言語文書データの生成が可能か否かを判定する(S702)。本実施の形態では、前述したように各目標言語毎に個別の目標言語文書生成プログラムを持つので、S701にて受信した要求言語情報に対応した目標言語文書生成プログラムが目標言語文書生成プログラム記憶部302bに記憶されているか否かで、S705の判定を行う。S705の判定の結果、目標言語文書データの生成が可能であれば、S706を行い、不可能であればS709に処理を移す。S706では、受信した中間形式文書データから目標言語文書データを生成する、目標言語文書生成処理を行う。S706の処理の詳細は後述する。次に、生成された目標言語文書データを送信文書メモリ303cに記憶し、送信文書とする(S710)。

【0060】S709では、要求言語情報に対応した目標言語文書が生成できない旨のエラー情報を生成して、そのエラー情報を送信文書メモリ303cに記憶し、これを送信文書とする。S707ではサーバー装置から受信し、受信文書メモリ303bに記憶した文書データを、送信文書メモリ303cに記憶させ、送信文書とする。

【0061】最後に送信文書メモリ303cに記憶された文書データをクライアント装置103に送信し(S710)、変換装置の処理を終了する。

【0062】なお、クライアント装置103が要求言語情報を送信しない構成であれば、変換装置103はS701にお

いて要求言語情報を受信しない構成としてもよい。この場合は、予め変換装置104に設定されている言語情報を要求言語情報として用いることとする。

【0063】なお、S702において要求言語情報の送信は行わない構成としてもよい。この場合は、例えばクライアント装置から図6(b)を受信したとすると、図6(d)で例示される情報をサーバー装置102に送信する。

【0064】次に、サーバー装置102の基本的な処理の流れを図8を用いて説明する。先ず、変換装置104からS703にて送信された要求文書の識別情報と、要求言語情報とを受信する(S801)。次に、受信した要求文書の識別情報と、要求言語情報とに対応する原文書データが原文書メモリ203cに記憶されているか否かを検索し、記憶されていればS806へ、記憶されていなければS803に処理を進める。S803では、受信した要求文書の識別情報に対応した中間形式文書データが中間形式文書データ記憶部203dに記憶されているか否かを判定し、記憶されていればS805に、記憶されていなければS804に処理を進める。

【0065】S805では、受信した要求文書の識別情報に対応した中間形式文書データを中間形式文書データ記憶部203dから読み出し、これを送信文書として送信文書メモリ203dに書き込む。S804では、受信した要求文書の識別情報に対応した文書データがない旨のエラー情報を作成し、それを送信文書として送信文書メモリ203dに書き込む。S806では、受信した要求文書の識別情報と、要求言語情報とに対応する原文書データを原文書メモリ203cから読みだし、送信文書として送信文書メモリ203dに書き込む。次に、送信文書メモリ203dに書き込まれた送信文書を変換装置104に送信して処理を終了する。

【0066】なお、変換装置104が要求言語情報をサーバー装置102に要求言語情報を送信しない構成の場合は、サーバー装置においても、S802及びS806の処理を持たない構成としてもよい。

【0067】本実施の形態においては、中間形式文書データ記憶部203dで記憶される中間形式文書データや、原文書メモリ203cに記憶される原文書データは、それぞれの文書データが1つのファイルに対応するものとし、文書データを指定する文書識別情報としては、中間形式文書データあるいは原文書データを収納しているファイル名を用いるものとする。なお、対応する中間形式文書データと原文書データとは、同一のファイル名を用い、記憶場所の違いにより区別する。この場合は、文書識別情報で示されたファイル名のファイルが存在するか否かでS802やS803の判定を行い、S805では文書管理プログラムがそのファイルのデータを読み込むこととなる。なお、文書識別情報としてファイル名を用いる方法以外に、文書管理プログラムとしてデータベース管理プログラムを用い、中間形式文書データ記憶部203dとしてデータベースの記憶領域を対応させ、文書識別情報としてデータベース内の文書データを特定可能な指示情報を用いても、よ

い。

【0068】次に、請求項2記載の多言語対応システムの変換装置が行う中間形式文書データの作成の手順を、図9、図10、図11を用いて説明する。まず、サーバー装置102より、所定の自然言語で記述された文書データである原文書データを受信し、一旦受信文書メモリ303bに保存する(S901)。原文書データは、先ず、形態素解析処理を受ける(S902)。S902では、翻訳辞書記憶部302dに記憶された形態素解析辞書を参照し、原文書を形態素に分割し、形態素列に変換し、各形態素の品詞の判定を10 行う。日本語から中間形式文書データを生成する一例として、図10(a)に示す原文書であれば、形態素解析処理により図10(b)のような形態素列からなる形態素解析結果が得られる。

【0069】次に、構文解析処理を行う(S903)。S903では、S902で解析された品詞情報を含む形態素列に、翻訳辞書記憶部302dに記憶された構文規則データと、構文解析辞書の情報を適用し、各形態素間の修飾関係を解析すると共に、格の意味的な役割を決定し、構文解析木を形成する。例えば、図10(b)のような形態素解析結果は、図10(c)で示される構文解析木からなる構文解析結果を得る。ここで、「S」は文を、「NP」は名詞句を、「ADV」は副詞を、「VP」は動詞句をそれぞれ表す。次に、依存構造解析処理を行う(S904)。S904では、S903の解析結果である構文解析木に基づき依存構造を形成する。図10(c)の構文解析結果に対して依存構造解析処理を行うと、図11(a)のような依存構造が得られる。ここで、「PRED」は述部を、「NOM」は依存構造名詞句を、「ADV」は副詞をそれぞれ表す。次に、S904で得られた依存構造のデータを中間形式文書データとして30 送信文書メモリ303cに一旦記憶した後、サーバー装置102に送信する(S905)。

【0070】本実施の形態では、中間形式表現として文の依存構造の表現を用い、前記のように文章を解析した結果得られる、依存構造で表現された文章の意味内容を表すデータを、中間形式文書データとして用いることとする。具体的には、図12(a)に記載するような構造要素1205の集合として記憶する。各構造要素1205は構造種別1201と、格情報1202と、辞書ポインター1203と、従属構造数1204の4つの項目を持つ。構造種別1201は述部とか、名詞句とかの種別をフラグとして記憶する。格情報1202は主格とか目的格、あるいは状態等の格の種別情報を必要に応じてフラグとして記憶する。辞書ポインター1203は、その構造要素に対応した単語の情報を収納している辞書データへのポインターを記憶する。従属構造数1204は、その構造要素に直接従属する構造要素がいくつあるかを記憶する。なお、従属構造数1204は、その構造要素に従属する構造要素にさらに従属する、つまり、入れ子として間接的に従属する要素の個数は含まない。

【0071】例えば 図11(a)に示された依存構造を

記憶した場合は、図12(b)のようなデータとなる。ここで、「K1」は主格を、「K2」は目的格を、「K3」は状態の格を表す。また、辞書ポインターは、図12(c)に例示された辞書データの番号1206を表している。辞書データには、原文書で用いられた言語の単語1207と、その品詞1208、及び他の言語の対応する単語1209や、さらに、各種の文法情報などが記憶される。図12(c)では、他の言語が英語である場合の例示である。なお、中間形式文書データとしては上記の方法以外に、人工的に定義された可読性のある言語(以下、共通言語という)を、中間形式表現として用い、その言語で記述された文書データを中間形式文書データとして用いてもよい。例えば、特開平6-325080公報に開示されているような、人工的に設計した言語で記述した文章を中間形式文書データとしてもよい。

【0072】次に、変換装置104が行うS706の目標言語文書生成処理の詳細を、図13と図11とを用いて説明する。目標言語文書生成処理は、本発明の目標言語文書生成手段の主体をなす目標言語文書生成プログラムが行う処理である。先ず、S704にて受信文書メモリ303bに記憶された中間形式文書データに対し、構文生成処理を行う(S1301)。構文生成処理では、中間形式文書データに記述された依存構造に基づいて、生成規則データを参照しつつ目標言語の構文木を生成し、語順を決定する。次に、形態素生成処理を行い(S1302)、単語の語尾等を必要に応じて変化させ、生成文書データの生成を完了する。目標言語が英語である一例を挙げれば、図11(a)で示される依存構造の文に対し、構文生成処理と、形態素生成処理の処理を行うと、図11(b)に示されるような目標言語文書生成結果が得られることとなる。ここで、「PRON」は代名詞、「V」は動詞、「ART」は冠詞、「N」は名詞をそれぞれ表す。

【0073】なお、本実施の形態においては、中間形式文書データとして記憶されている依存構造データをそのまま構文生成処理への入力としたが、構文生成処理への入力の前に依存構造変換処理を行ってもよい。これは具体的には、中間形式文書データに記憶されている文章の依存構造をそのまま用いるのではなく、中間形式変換手段に対応する依存構造変換処理を行い、目標言語に対応した依存構造に変換してから、構文生成処理への入力とする。

【0074】なお、本実施の形態に関わる前記の形態素解析処理、構文解析処理、依存構造解析処理、構文生成処理、及び形態素生成処理のそれぞれは、電子技術総合研究所と京都大学による「日英科学技術文献の速報システムに関する研究一言語処理システムの開発に関する報告書」や、特開昭62-203273号公報に開示されている公知の技術である。

【0075】

【発明の効果】以上説明したことから明らかなように

本発明の請求項1に記載の多言語対応通信システムによれば、クライアントから所望の記述言語を指定でき、それに従って中間形式文書データの形式で記憶された文書データから、変換装置が自然言語で記述された文書データを生成し、クライアント装置で出力できるので、サーバー装置に多数の言語で記述した文書データを保存したり、クライアント装置やサーバー装置に言語翻訳プログラムを搭載することなく簡易に構成でき、所望の言語で表記された文書データの閲覧ができる。

【0076】また、請求項2に記載の多言語対応通信システムは、自然言語で記述された原文書データを、サーバー装置から変換装置に送信し、変換装置で原文書データから中間形式文書データを生成し、生成された中間形式文書データをサーバー装置に送り返し、サーバー装置でその中間形式文書データを記憶することで、中間言語の作成に要するプログラムや辞書データをサーバー装置毎に用意する必要がなくなる。

【0077】また、請求項3に記載の多言語対応通信システムは、例えば従来用いられているサーバー装置のような、中間形式文書データが保存されていないサーバー装置や、自然言語で記述された文書データを出力するサーバー装置等に対応することができ、文書データが得られないという可能性を減少させることが出来る。

【0078】また、請求項4に記載の多言語対応通信システムは、変換装置において複数の言語に対応する文書データを生成することが可能となり、各言語毎に異なった変換装置を用意する必要がなくなる。

【0079】また、請求項5に記載の多言語対応通信システムは、所望の言語による文書データがサーバー装置に保存されている場合は、その保存されている文書データを、変換装置を経由してクライアント装置に送信することが可能となる。よって、必要な場合のみ中間形式文書データから所望の言語で記述された文書データを行えばよく、処理の高速化を行うことが出来る。

【図面の簡単な説明】

【図1】本発明の実施の形態の多言語対応通信システムの概要を示す説明図である。

【図2】本実施の形態のサーバー装置のブロック図である。

【図3】本実施の形態の変換装置のブロック図である。

【図4】本実施の形態のクライアント装置のブロック図である。

【図5】本実施の形態のクライアント装置の処理の流れを示すフローチャートである。

【図6】本実施の形態の文書識別情報と要求言語情報との説明図である。

【図7】本実施の形態の変換装置の処理の流れを示すフローチャートである。

【図8】本実施の形態のサーバー装置の処理の流れを示すフローチャートである。

【図9】本実施の形態の中間形式文書データの作成に関する処理の流れを示すフローチャートである。

【図10】本実施の形態の中間形式文書データを作成する処理の説明に供する説明図である。

【図11】本実施の形態の中間形式文書データ作成の処理と、目標言語文書生成処理の説明に供する説明図である。

【図12】本実施の形態の中間形式文書データの説明に供する説明図である。

【図13】本実施の形態の目標言語文書生成処理のフローチャートである。

【図14】従来の通信システムの概要を示す説明図である。

【符号の説明】

101 通信ネットワーク

102 サーバー装置

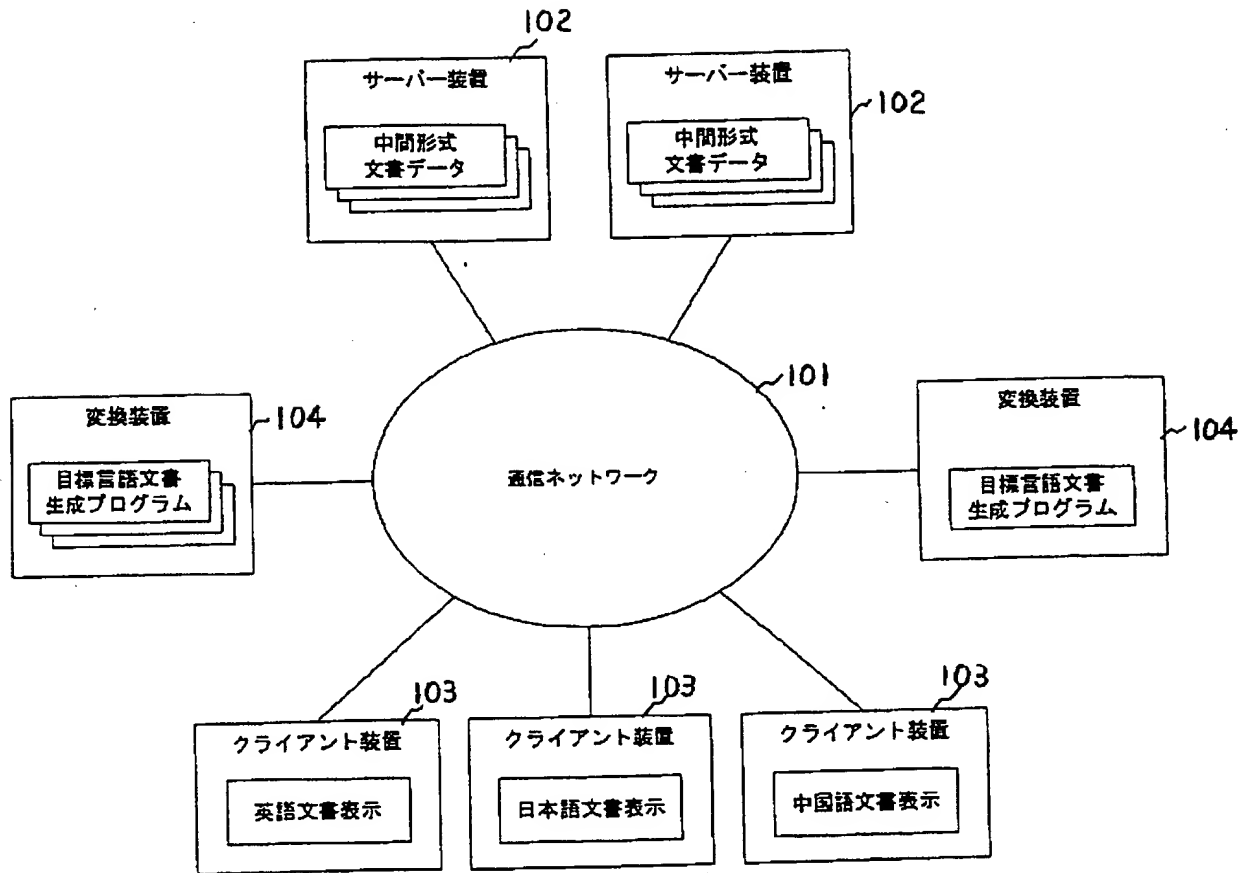
103 クライアント装置

104 変換装置

203d 中間形式文書データ記憶部

302b 目標言語生成プログラム記憶部

【図1】



【図6】

```
GET <ロケーション情報> <バージョン情報><CRLF>
<各種ヘッダー情報>
Accept-Language:<要求言語><CRLF>
<CRLF>
```

(a)

```
GET http://server1/pub/sample.html HTTP/1.0<CRLF>
Accept-Language:jp<CRLF>
<CRLF>
```

(b)

```
GET /pub/sample.html HTTP/1.0<CRLF>
Accept-Language:jp<CRLF>
<CRLF>
```

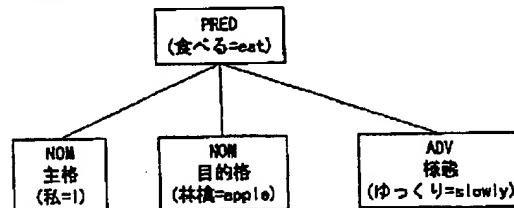
(c)

```
GET /pub/sample.html HTTP/1.0<CRLF>
<CRLF>
```

(d)

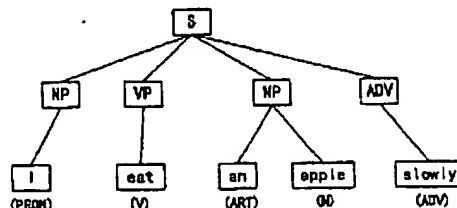
【図11】

依存構造



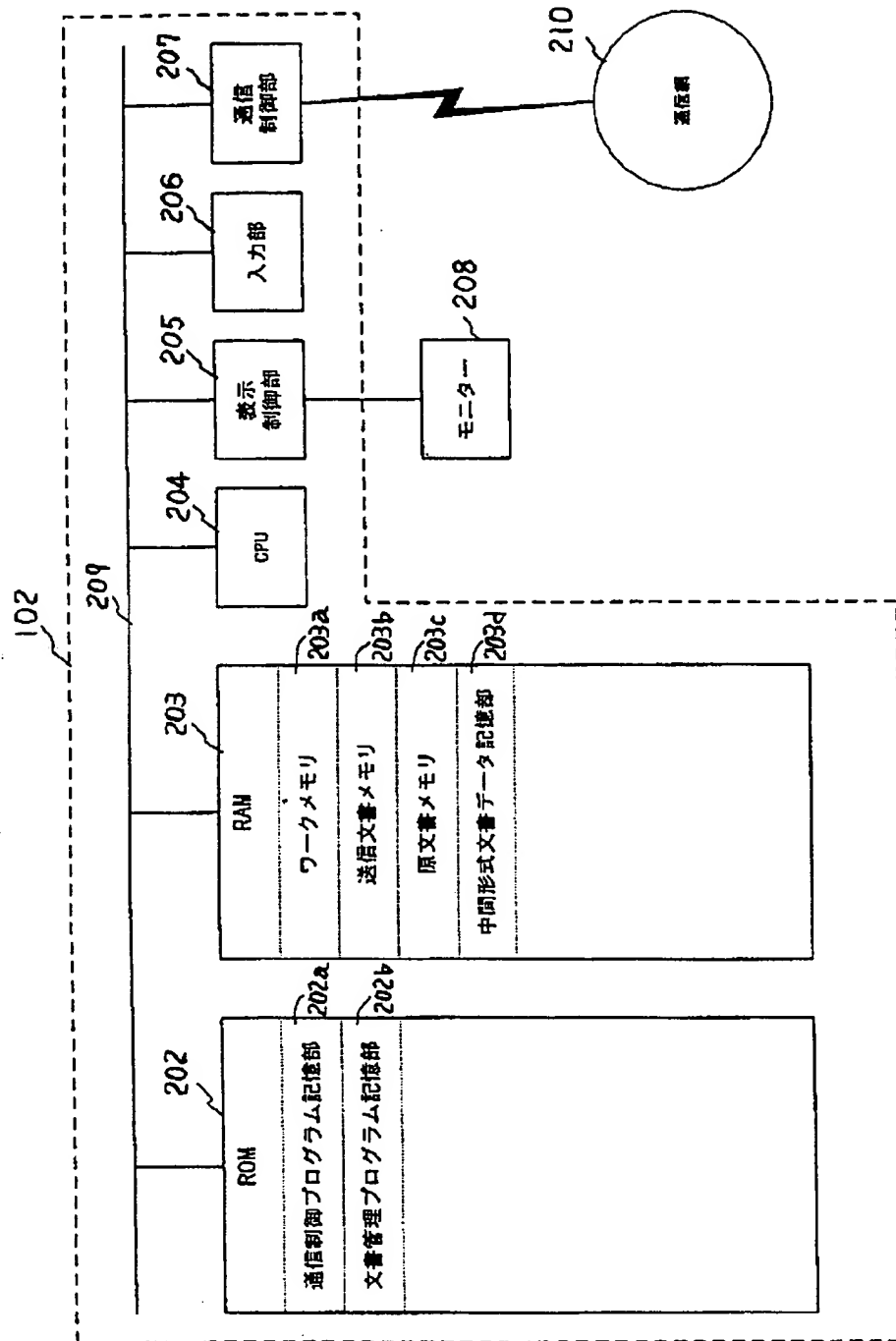
(a)

目標言語生成結果

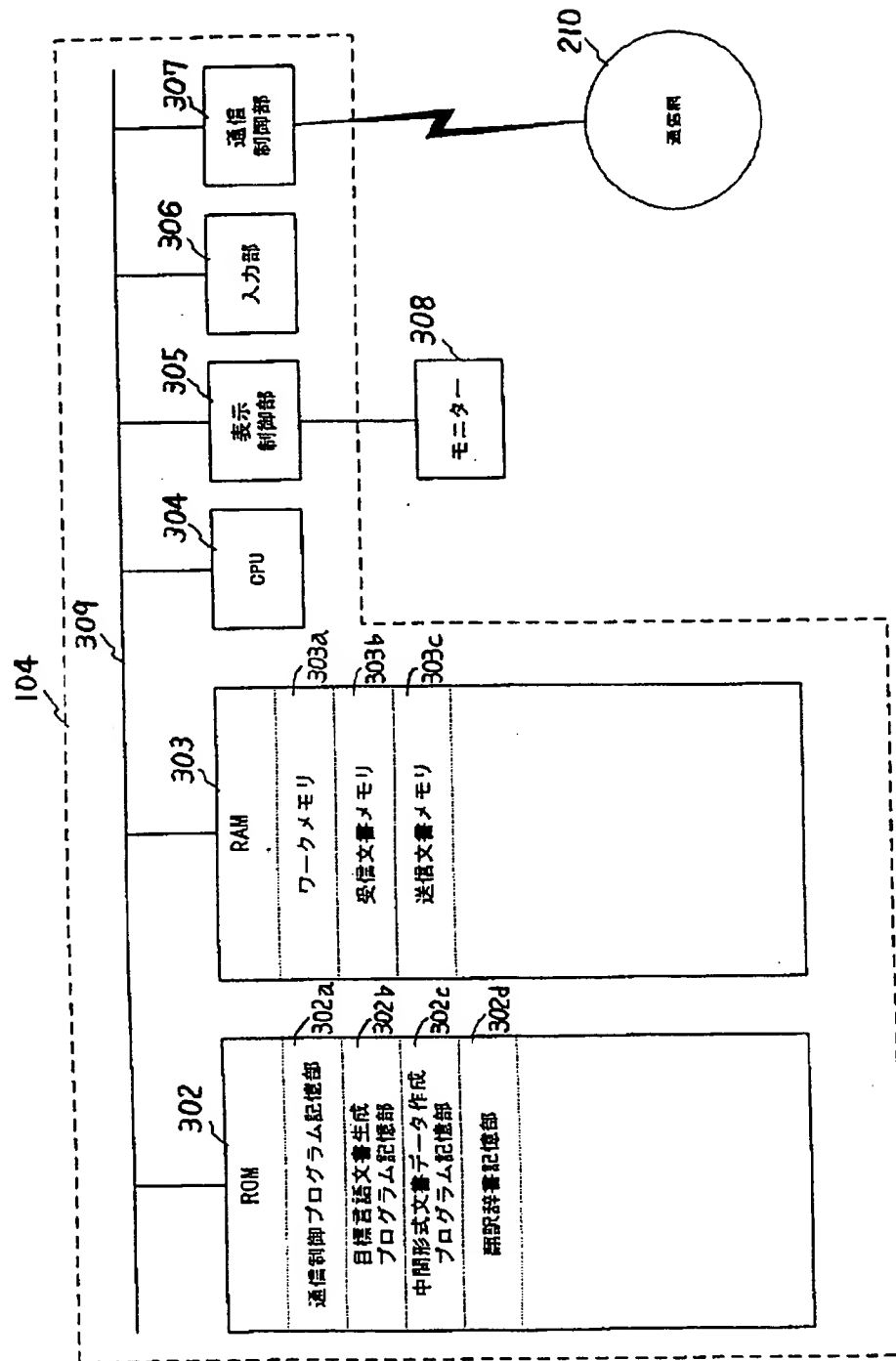


(b)

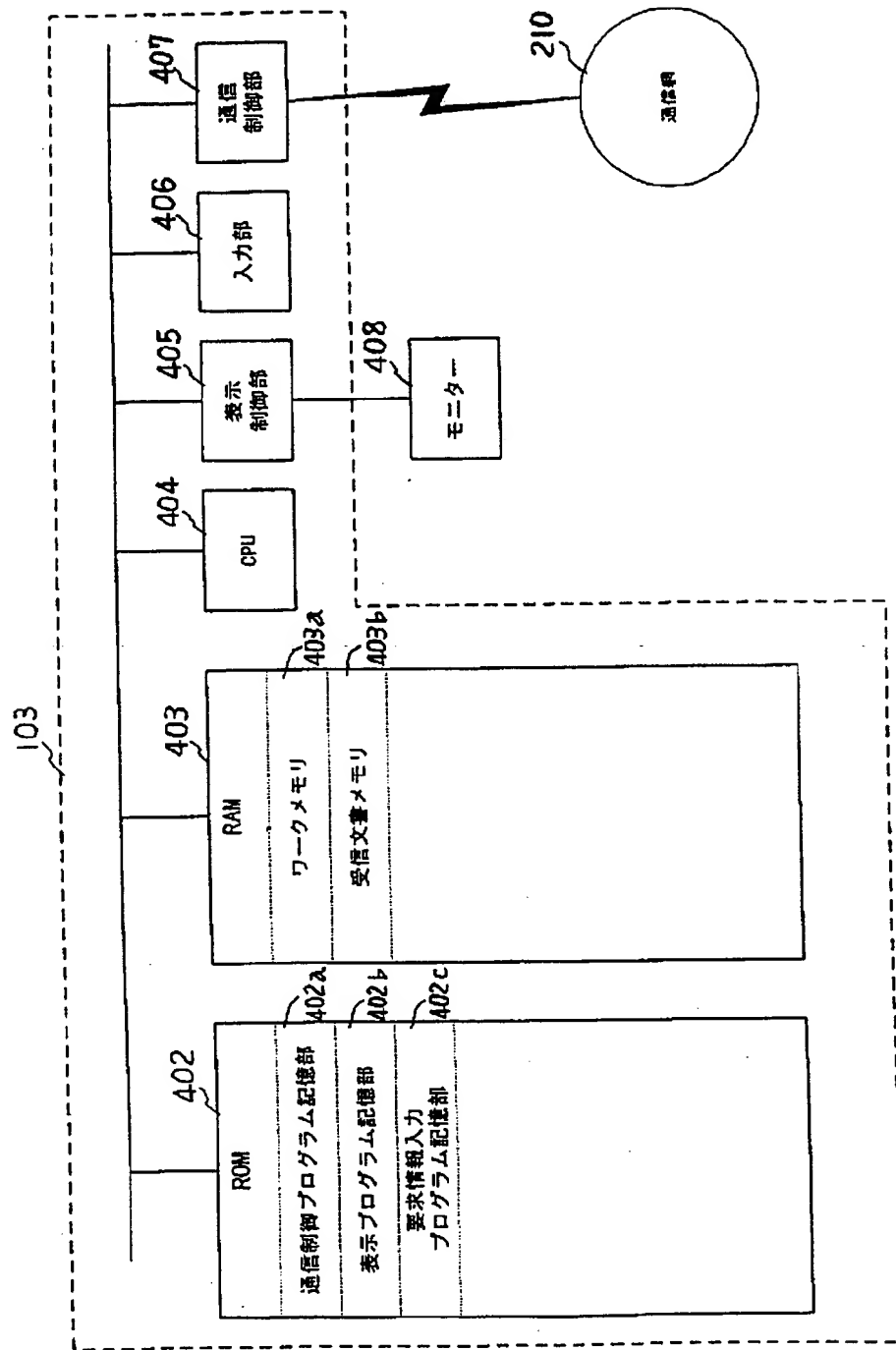
【図2】



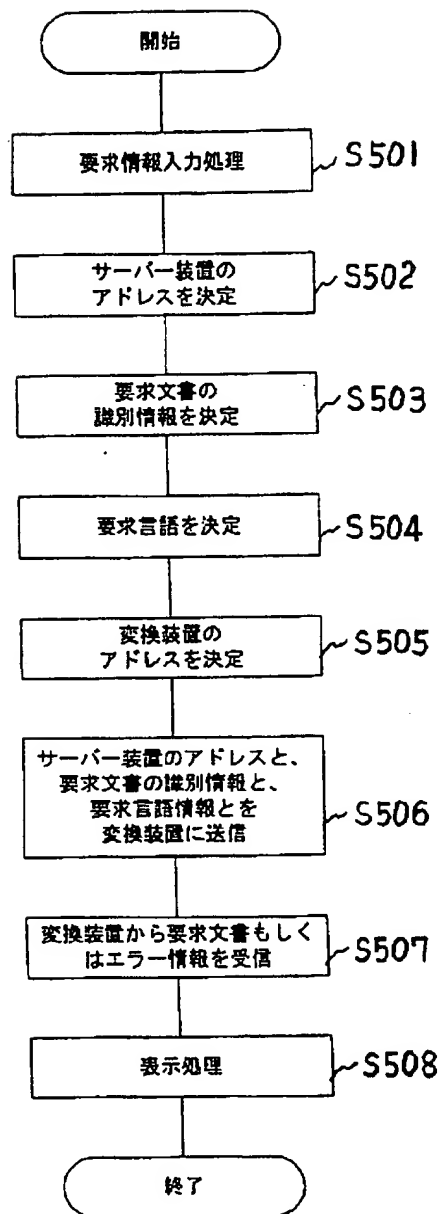
【図3】



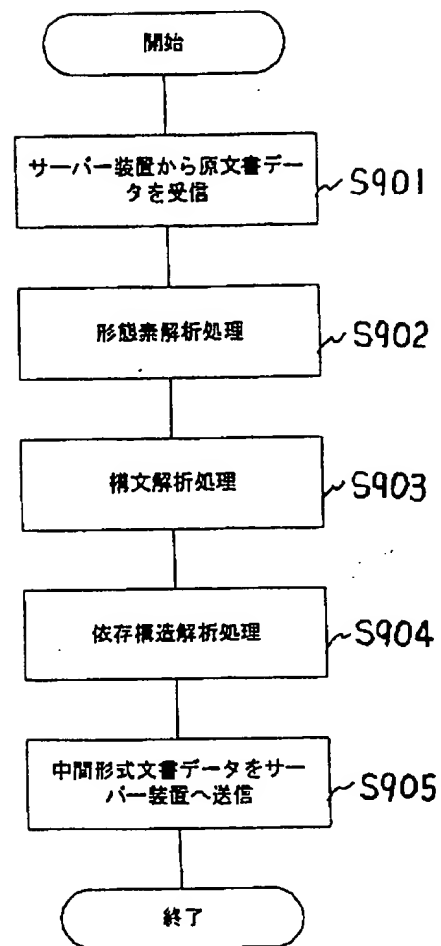
【図4】



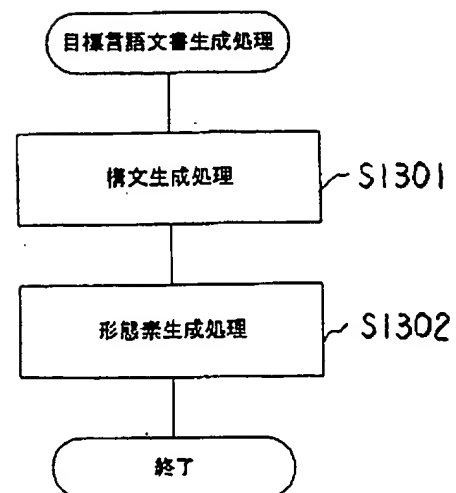
【図5】



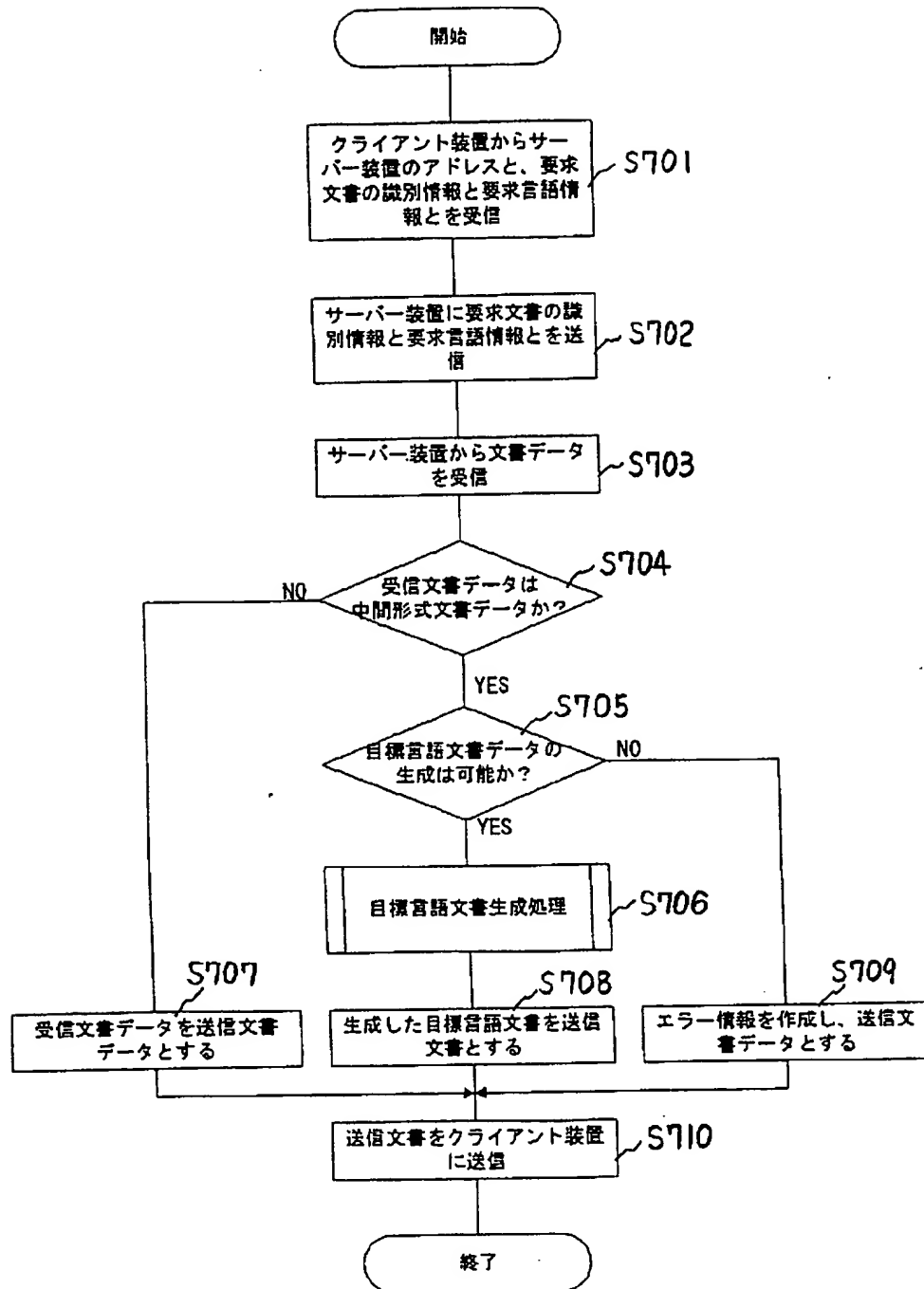
【図9】



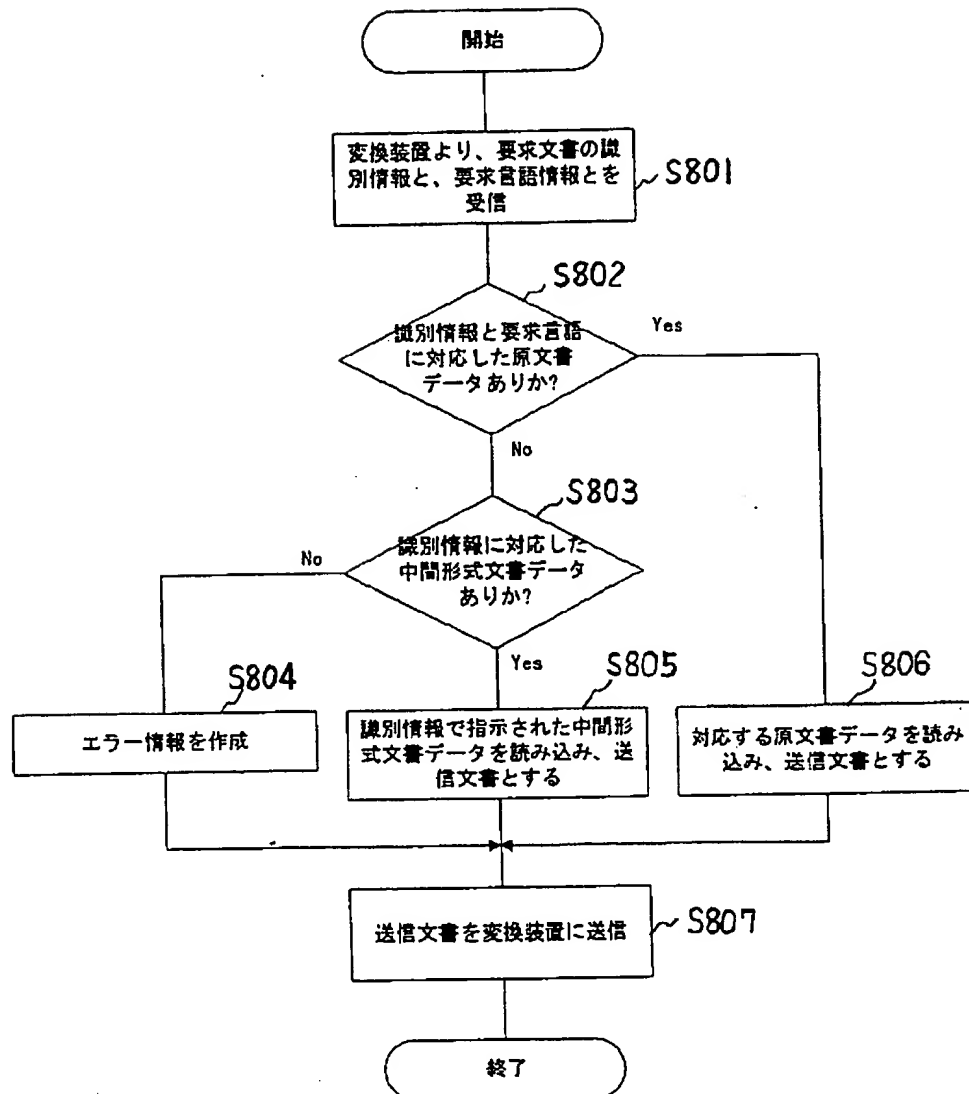
【図13】



【図7】



【図8】



【図10】

原文:
私は林檎をゆっくり食べる

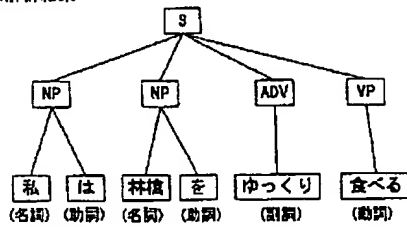
(a)

形態素解析結果

私 は 林檎 を ゆっくり 食べる
(名詞) (助詞) (名詞) (助詞) (副詞) (動詞)

(b)

構文解析結果



(c)

【図12】

1201	1202	1203	1204	
構造種別	格情報	辞書	従属構造数	~1205
構造種別	格情報	辞書	従属構造数	

(a)

PRED	-	001	3
NOM	K1	002	0
NOM	K2	003	0
ADV	K3	004	0

(b)

1206	1207	1208	1209	
001	食べる	動詞	eat	- - -
002	私	代名詞	I	- - -
003	林檎	名詞	apple	- - -
004	ゆっくり	副詞	slowly	- - -

(c)

【図14】

